

06/29/00
jc832 U.S. PTO

07-02-00

A

Please type a plus sign (+) inside this box [+]

PTO/SB/05 (12/97)

Approved for use through 09/30/00. OMB 0651-0032

Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

UTILITY PATENT APPLICATION TRANSMITTAL
(Only for new nonprovisional applications under 37 CFR 1.53(b))

Attorney Docket No. 81862.P183

Total Pages 2

First Named Inventor or Application Identifier Bruce Wilford

Express Mail Label No. EL627463900US

jc832 U.S. PTO

09/607584

06/29/00

ADDRESS TO: Assistant Commissioner for Patents
Box Patent Application
Washington, D. C. 20231

APPLICATION ELEMENTS

See MPEP chapter 600 concerning utility patent application contents.

1. X Fee Transmittal Form
(Submit an original, and a duplicate for fee processing)
2. X Specification (Total Pages 32)
(preferred arrangement set forth below)
 - Descriptive Title of the Invention
 - Cross References to Related Applications
 - Statement Regarding Fed sponsored R & D
 - Reference to Microfiche Appendix
 - Background of the Invention
 - Brief Summary of the Invention
 - Brief Description of the Drawings (if filed)
 - Detailed Description
 - Claims
 - Abstract of the Disclosure
3. X Drawings(s) (35 USC 113) (Total Sheets 4)
4. Oath or Declaration (Total Pages)
 - a. Newly Executed (Original or Copy)
 - b. Copy from a Prior Application (37 CFR 1.63(d))
(for Continuation/Divisional with Box 17 completed) (**Note Box 5 below**)
 - i. DELETIONS OF INVENTOR(S) Signed statement attached deleting inventor(s) named in the prior application, see 37 CFR 1.63(d)(2) and 1.33(b).
5. Incorporation By Reference (useable if Box 4b is checked)
The entire disclosure of the prior application, from which a copy of the oath or declaration is supplied under Box 4b, is considered as being part of the disclosure of the accompanying application and is hereby incorporated by reference therein.
6. Microfiche Computer Program (Appendix)

09607584-062900

7. _____ Nucleotide and/or Amino Acid Sequence Submission
(if applicable, all necessary)
a. _____ Computer Readable Copy
b. _____ Paper Copy (identical to computer copy)
c. _____ Statement verifying identity of above copies

ACCOMPANYING APPLICATION PARTS

8. _____ Assignment Papers (cover sheet & documents(s))
9. _____ a. 37 CFR 3.73(b) Statement (where there is an assignee)
_____ b. Power of Attorney
10. _____ English Translation Document (if applicable)
11. _____ a. Information Disclosure Statement (IDS)/PTO-1449
_____ b. Copies of IDS Citations
12. _____ Preliminary Amendment
13. X Return Receipt Postcard (MPEP 503) (Should be specifically itemized)
14. _____ a. Small Entity Statement(s)
_____ b. Statement filed in prior application, Status still proper and desired
15. _____ Certified Copy of Priority Document(s) (if foreign priority is claimed)
16. X Other: Certificate of Express Mail with copy of postcard showing contents of
Express Mail package.

17. **If a CONTINUING APPLICATION**, check appropriate box and supply the requisite information:
____ Continuation ____ Divisional ____ Continuation-in-part (CIP)
of prior application No: _____

18. Correspondence Address

____ Customer Number or Bar Code Label _____
(Insert Customer No. or Attach Bar Code Label here)

or

X Correspondence Address Below

NAME Sang Hui Michael Kim Sang Hui Kim
BLAKELY, SOKOLOFF, TAYLOR & ZAFMAN LLP

ADDRESS 12400 Wilshire Boulevard
Seventh Floor

CITY Los Angeles STATE California ZIP CODE 90025-1026

Country U.S.A. TELEPHONE (408) 720-8598 FAX (408) 720-9397

UNITED STATES PATENT APPLICATION

FOR

METHOD AND APPARATUS PROVIDING IMPROVED STATISTICS
COLLECTION FOR HIGH BANDWIDTH INTERFACES
SUPPORTING MULTIPLE CONNECTIONS

Inventors:

Bruce Wilford
Jan Medved
Stephan Crandall

Prepared by:

BLAKELY, SOKOLOFF, TAYLOR & ZAFMAN
12400 WILSHIRE BOULEVARD
SEVENTH FLOOR
LOS ANGELES, CALIFORNIA 90025
(408) 720-8300

"Express Mail" mailing label number:

Date of Deposit:

I hereby certify that I am causing this paper or fee to be
deposited with the United States Postal Service "Express Mail
Post Office to Addressee" service on the date indicated above
and that this paper or fee has been addressed to the
Commissioner of Patents and Trademarks, Washington, D. C.
20231

(Typed or printed name of person mailing paper or fee)

(Signature of person mailing paper or fee)

(Date signed)

EL627463900US

6/29/00

JUANITA BRUSCE

Juanita Bruce

6/29/00

006290"48520960

**METHOD AND APPARATUS PROVIDING IMPROVED STATISTICS
COLLECTION FOR HIGH BANDWIDTH INTERFACES
SUPPORTING MULTIPLE CONNECTIONS**

5

FIELD OF THE INVENTION

The present invention relates generally to digital communications. More particularly, the present invention relates to network switches and interfaces.

Specifically, the present invention relates to a method and apparatus providing
10 improved statistics collection for high bandwidth interfaces supporting multiple connections.

BACKGROUND OF THE INVENTION

A common type of network device is a network switch. A network switch forwards traffic or data units (e.g., cells or packets) from a source network
15 component or components to a destination network component or components. Typically, a network switch includes interfaces such as, for example, line cards, which are used to process network traffic and to pass selectively the traffic in and out of the switch. A line card is capable of supporting multiple high bandwidth channels or connections on bi-directional ports.

20 A common operation performed by a line card is accumulating and collecting statistics about a connection. Accumulating and collecting statistics about a connection has many useful applications. For example, if a connection is for a telephone call, determining the amount or types of cells or packets passing via the connection during a specific time period is important for billing purposes.

Thus, as traffic passes through a line card, statistics are counted and collected. Current line cards count statistics for each connection using counters. For example, a traffic cell passing through a line card may increment a traffic cell counter for a connection by one and the incremented count is stored in a
5 memory.

Consequently, as line card speeds and number of connections per line cards increase, a limitation with current network interfaces is that a larger number of statistics have to be collected in shorter periods of time. A common method of collecting statistics for connections is polling for stored counted
10 statistics in set time intervals for each connection regardless of the amount or type of traffic passing in each connection.

A disadvantage with the above polling method is data loss. That is, if a connection having high traffic must wait a certain time period before being polled, a counter for the high traffic connection may roll over thus providing
15 unreliable and inaccurate statistical information. Furthermore, polling statistical information for a connection having a small amount or no traffic wastes processing time.

Therefore, avoiding data loss and reducing wasted processing time is critical in maintaining reliable and accurate statistical information for multiple
20 connections supported by a network switch or interface.

SUMMARY OF THE INVENTION

A method and apparatus are disclosed that provide improved statistics collection for high bandwidth interfaces supporting multiple connections.

Statistics are counted for a plurality of connections. A determination is made to
5 determine if a count value for each of the connections is greater than or equal to a threshold. Statistics are collected for connections having a determined count value greater than or equal to threshold before connections having a count value below the threshold.

Other features and advantages of the present invention will be apparent
10 from the accompanying drawings, and from the detailed description, which follows below.

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention is illustrated by way of example and not limited by the figures of the accompanying drawings in which like references indicate similar elements and in which:

5 **Figure 1** is an illustration of an exemplary network switch for practicing the invention;

Figure 2 is an internal functional block diagram of a line card according to one embodiment;

Figure 3 is an illustration of an exemplary stats collection architecture;

10 **Figure 4A** is a flow chart illustrating an exemplary high priority interrupt operation;

Figure 4B is a flow chart illustrating an exemplary improved polling operation; and

15 **Figure 5** is a flow chart illustrating an exemplary operation for collecting statistics.

DETAILED DESCRIPTION

A method and apparatus are described that provide improved statistics collection for high bandwidth interfaces supporting multiple connections.

Statistics are counted for a plurality of connections. A determination is made to
5 determine if a count value for each of the connections is greater than or equal to a threshold. Statistics are collected for connections having a determined count value greater than or equal to the threshold before connections having a count value below the threshold.

By collecting statistics for connections having respective count values
10 greater than or equal to a threshold before collecting statistics for connections having respective count values below the threshold, loss of statistical information for high traffic connections can be avoided. That is, statistical information for high traffic connections is collected before a counter associated with a respective high traffic connection rolls over. Furthermore, processing load is reduced and
15 used efficiently by collecting statistics at a rate necessary for low traffic connections.

The apparatus and method described herein can thus avoid data loss and reduce wasted processing time in maintaining reliable and accurate statistics for multiple connection supported by a network switch or interface.

20 In the following embodiments, a network switch and interface are presented in the context of collecting statistical information for traffic cells. For example, the network interface may be a line card supporting multiple

connections, which communicate traffic cells for an asynchronous transfer mode (ATM) switch. Nevertheless, other types of data units for varying types of network switches may be used such as, for example, data frames, blocks or packets.

5 Furthermore, in the following embodiments, statistical information is referred to as "stats." In addition, statistical information collection is referred to as "stats collection."

10 **Figure 1** is an illustration of an exemplary network 100 switch for practicing the invention. Alternatively, other types of network devices such as, for example, a network router, may be used for practicing the invention.

Referring to **Figure 1**, network switch 100 includes a plurality of line cards 102 (1 through N), which may support a plurality of connections on multiple ports. Network switch 100 is a network device that forwards selectively traffic cells on connections supported by line cards 102. Network switch 100 may support a number of protocols. For example, network switch 100 may be an asynchronous transfer mode (ATM) switch supporting multiple virtual connections ("VCs"). Alternatively, network switch 100 may be an Ethernet or Internet switch.

20 Line cards 102 are interfaces for network switch 100 that receive and forward traffic cells to and from network switch 100. Line cards 102 may be inserted or removed from the network switch 100. Line cards 102 may also support a number of different types of network ports. For example, line cards

102 may support optical carrier (OC) bi-directional network ports. Alternatively, line cards 102 may support other types of ports such as, for example, synchronous transport system (STS) ports or synchronous digital hierarchy (SDH) ports.

5 Furthermore, each line card may have any number of network ports operating at different speeds and supporting a plurality of connections. In the example of **Figure 1**, each line card is shown to be a 10 Gigabit/s line card having four 2.5 Gigabit/s optical carrier OC48 bi-directional network ports. The four (OC48) network ports may support up to one million connections.

10 **Figure 2** is an internal functional block diagram of a line card for line cards 102 according to one embodiment. For purposes of explanation, the line card of **Figure 2** is referred to as line card 102. Line card 102 may represent any of the line cards 1 through N. For purposes of clarity, line card 102 is shown as uni-directional, however, line card 102 may be bi-directional. Furthermore, basic
15 internal circuitry and components of line card 102 are not described in detail as not to obscure invention.

Referring to **Figure 2**, line card 102 includes a service module 204 having a statistical ("stats") collection module 206. Service module 204 may be hardware or software or a combination of both that processes connection traffic of
20 incoming cells ("cells") for network switch 100. Service module 204 operates to set up connections for network switch 100. For example, as cells are being received by line card 102, service module 204 determines which network ports

on which line cards to forward the cells to. Service module 204 processes a header field in the cells to determine which network ports to forward the cells to. Service module 204 forwards the cells to network switch 100 such that the cells are switched to desired network ports on desired line cards.

5 Stats collection module 206 may also be hardware or software or a combination of both that processes cells to accumulate and collect statistical information for connections based on the amount or type of cells being communicated via the connections. For example, **Table 1** shown below lists exemplary types of statistical information that may be accumulated and collected
10 by stats collection module 206 for line card 102.

Table 1
(Exemplary Statistical Information)

Number of Cells per Virtual Connection (VC)
Number of Cells with the Cell Loss Priority Bit set to 0 (CLP0)
Number of Cells with the Cell Loss Priority Bit set to 1 (CLP1)
Number of CLP1 cells dropped
Number of Operation and Management (OAM) cells

15 The exemplary statistical information shown in **Table 1** may be accumulated and collected for each connection supported by line card 102. Stats collection module 206 may include a plurality of adders wherein each adder is associated with a connection. Furthermore, an adder may be associated for each

stat of each connection. For example, an adder may correspond to a connection to count the number of cells that are being communicated via the connection. The adder is then incremented for each cell passing through line card 102 for that connection. Thus, adders in stats collection module 206 are continuously counting stats for each received cell.

Exemplary Stats Collection Architecture

Figure 3 is an illustration of an exemplary stats collection architecture 206 for the stats collection module of line card 102. The exemplary stats collection architecture 206 illustrates the internal components and operations of line card 102, which provide improved statistics collection for network interfaces supporting multiple connections.

For purposes explanation, the following description of exemplary stats collection architecture 206 refers to a single adder for a single stat. However, exemplary stats collection architecture 206 may include a plurality of adders counting stats for a plurality traffic types and connections.

Referring to **Figure 3**, the exemplary stats collection architecture 206 includes a translation table 326 and cell classifier 316, which process incoming cells 320, memory address generator 340 generating memory addresses, stats memory 310 storing stats for connections, adder 314 counting stats based on addresses from address generator 340, comparator 350 comparing count values with threshold values, memory 360 storing threshold values, first-in-first-out (FIFO) control unit 308 generating interrupt signals and controlling a FIFO buffer

304 storing logical connection identifiers (LCIs), central processing unit (CPU)

302 collecting stats, and CPU memory 318 storing collected stats.

The interconnections for exemplary stats collection architecture 206 will now be described in detail. Translation table 326 and cell classifier 316 are

5 coupled with memory address generator 340, which is coupled to stats memory 310. Translation table 326 is also coupled with FIFO control unit 308 via FIFO buffer 304. Adder 314 is coupled with stats memory 310 and comparator 350.

Comparator 350 is coupled with memory 360 that stores threshold values and

FIFO control unit 308. FIFO control unit 308 is also coupled with CPU 302 via

10 interrupt signal line 306 and FIFO buffer 304. CPU 302 is also coupled with stats memory 310 (interconnection not shown for clarity) and CPU memory 318.

The exemplary stats collection architecture 206 functions will now be described in detail. Translation table 326 is a processing unit that processes a header field 322 for each cell of incoming cells 320. Cell 320 includes a header

15 field 322 and data field 324. Translation table 326 may be a physical device or a software module providing instructions for a processing device. Although one cell 320 is depicted for incoming cells, the cell 320 may represent a stream of incoming cells. Header field 322 includes information such as, for example, cell or connection type information (i.e., VCI/VPI), destination information, or

20 switching information. Data field 324 includes information that is to be forwarded.

Cell classifier 316 is a processing unit that may process both header field 322 and data field 324 of cell 320. Cell classifier 316 may be a physical device or a software module providing instructions for a processing device. Cell classifier 316 operates to classify the "type of cell" for cell 320 based on the header field 322 or data field 324 or both. Cell classifier 316 sends the "type of cell" information to memory address generator 340. For example, if cell 320 is an OAM cell, cell classifier 316 classifies cell 320 as an OAM cell and informs memory address generator 340 that cell 320 is an OAM cell.

Translation table 326 operates to process the header field 322 of cell 320 and to translate the information in the header field 322 into a logical connection identifier. The logical connection identifier (LCI) is an internal representation to identify a specific connection and to a specific stream of cells for line card 102. CPU 302 can determine wherein in stats memory 310 to collect a specific stat for a specific "cell type" of a specific connection (LCI). That is, if cell 320 has an LCI and "cell type" corresponding to stat 312, then CPU 302 can collect all stats for a connection associated with that cell. Translation table 326 sends the LCI information to memory address generator 340. Translation table 326 may also send LCI information to FIFO buffer 304 (if instructed by FIFO control unit 308).

Memory address generator 340 is a processing unit that processes LCI information from translation table 326 and "type of cell" information from cell classifier 316. That is, memory address generator 340 can associate a memory address for every LCI and "type of cell" information for an incoming cell.

Memory address generator 340 may be a physical device or a software module providing instructions for a processing device. Memory address generator 340 operates to generate an address of a location in stats memory 310, which is used by adder 314 to count a stat 312 stored at the address. For example, if adder 314 is designated to count a stat 312 at a specific memory address for cell 320, memory address generator 340 will generate an address at location 312 in stats memory 310 for adder 314, which will use that address to increment a count stored in stat 312.

Stats memory 310 is a memory device, which operates to store a plurality of statistical information for each connection. For example, stats memory 310 may store statistical information as described in **Table 1** above. The statistical information may be accumulated or updated as each cell is being received by the line card 102. That is, as each cell is received, an adder corresponding to the stat for the cell is incremented by one. Furthermore, although a single stat memory is shown, the exemplary stats collection architecture 206 may utilize any number of memories for storing stats.

Adder 314 is a counting device. For one embodiment, adder 314 is a 32-bit counter. Adder 314 operates to increment by one a stat stored at the address generated by memory address generator 340. For purposes of clarity, one adder 314 is shown. However, exemplary stats collection architecture 206 may have a plurality of adders such that an adder is associated with each "type of cell" information and LCI. In the example of **Figure 3**, adder 314 is associated with

the LCI information and "type of cell" information related to stat 312 stored in stats memory 310. Stat 312 stores statistical information related to cell 320.

Adder 314 increments a count value represented by stat 312. To increment or accumulate stat 312, stat 312 is retrieved using the address from address generator 340 and inputted into adder 314. Adder 314 then increments the count value for stat 312 by one and stores the incremented count value as stat 312 back in the same location retrieved from in stats memory 310. For example, if stat 312 stores the total number of OAM cells for connection 1 (LCI 1) and cell 320 is an OAM cell for connection 1 (LCI 1), memory address generator 340 will generate an address for the location of stat 312 stored in stats memory 310. Assuming stat 312 stores a count value of "100," adder 314 will increment the count value of "100" by one to generate an incremented count value of "101," which is stored as stat 312 in stats memory 310 at the same location.

The count value of adder 314 is also inputted to comparator 350.

Comparator 350 is a device that operates to compare count values in adder 314 with programmable threshold values stored in memory 360. For purposes of clarity, one comparator 350 is shown. However, exemplary stats collection architecture 206 may have a plurality of comparators such that a comparator is associated with each adder. Comparator 350 compares count values for adder 314 with threshold values stored in memory 360. For example, comparator 350 can determine whether a count value for adder 314 is greater than or equal to a specific threshold value stored in memory 360. Comparator 350 operates to

inform FIFO control unit 308 if a count value of adder 314 is greater than or equal the threshold.

FIFO control unit 308 is a processing device that processes information from comparator 350. FIFO control unit 308 operates to store the LCI related to stat 312 that is above a threshold, into FIFO buffer 304. FIFO buffer 304 may store a plurality of LCIs (LCI_1 through LCI_N). Consequently, LCI_1 through LCI_N refer to high priority stats - stats having a corresponding adder with a count value that is greater than or equal to a threshold. As such, FIFO control unit 308 operates to send an interrupt signal to CPU 302 if the count value of adder 314 is greater than or equal to the threshold. The interrupt signals informs CPU 302 that stat 312 is greater than or equal to the threshold and that consequently all stats related to that connection (LCI) should be collected before collecting stats related to LCIs having all stats below the threshold.

CPU 302 is a processing unit that polls and collects statistical information by reading ("stats") in stats memory 310 for each connection of line card 102. CPU 302 may include its own counters or adders to count the polled or collected stats from stats memory 310. For one embodiment, CPU 302 has internal counters that are wider than adder 314 for stats collection module 206. For example, the internal counters for CPU 302 may be 64-bit counters and adder 314 may be a 32-bit counter. As such, CPU 302 may maintain a larger stat count than adder 314. CPU 302 may store polled and collected stats in CPU memory 318. CPU memory 318 may be located within CPU 302 or outside of CPU 302.

006290"48520960

For one embodiment, CPU 302 may poll each stat in stats memory 310 at regular intervals to collect the stats for each connection, but may poll stats for the connections out of order based on a received interrupt from FIFO control unit 308 via interrupt signal line 306. For example, in one implementation, CPU 302 collects stats based on adder 314 being below or greater than or equal to a threshold. That is, if CPU 302 was polling or collecting statistics for connections at regular time intervals and stat 312 is incremented by adder 314 to produce a result that is greater than or equal to a threshold, FIFO control unit 308 would send an interrupt to CPU 302. CPU 302 would then read LCI₁ from FIFO buffer 304 and retrieve all the stats related to LCI₁ from stats memory 310 before polling or collecting stats for other connections.

Thus, CPU 302 will poll stats for high traffic connections more quickly than for low traffic connections, which prevents collecting unreliable and inaccurate stats. As such, CPU 302 can poll at a rate necessary for low traffic connections

Exemplary Operations

The exemplary operations for the exemplary stats collection architecture 206 will now be described with reference to Figures 4A, 4B, and 5. Figures 4A and 4B operate in parallel, however, for purposes of explanation, Figures 4A and 4B will be described separately.

Figure 4A is a flow chart illustrating an exemplary high priority interrupt operation 400. For purposes of explanation, operation 400 refers to the

exemplary stats collection architecture 206 of **Figure 3**. For purposes of explanation, operation 400 refers to a single adder 314, single stat 312, and a single comparator 350. Operation 400, however, may apply to a plurality of adders, stats, and comparators operating concurrently. Also, for purposes of
5 explanation, operation 400 begins at operation 402.

At operation 402, cell classifier 316 classifies a "type of cell" for cell 320 and translation table 326 selects an LCI for cell 320. The LCI and "type of cell" are combined into a memory address by memory address generator 340. The memory address determines which stat 312 is to be counted in stats memory 310.
10 That is, cell 320 relates to stat 312 for which adder 314 increments a count value for stat 312.

At operation 404, adder 314 increments the count value for stat 312 by one for cell 320. Adder 314 then stores the incremented count value back in stats 312 at the same location in stats memory 310. For example, if stat 312 stored total
15 number of cells, adder 314 will then add one to the total number of cells and store the incremented count value at the same location of stat 312. The count value for adder 314 is also input to comparator 350, which determines if the count value for adder 314 is greater than or equal to a threshold.

At operation 406, comparator 350 compares the count value of adder 314
20 with a programmable threshold value stored in memory 360 to determine if the count value for adder 312 is greater than or equal to the threshold.

For example, adder 314 may be a 32-bit adder and may have a threshold value that is 50% of a maximum count value for 32-bit adder. Thus, a 32-bit adder, will count to 50% its maximum count value when its count value is greater than 0x80000000.

- 5 Comparator 350 can be configured or programmed to determine if adder 314 has reached or is above of a count value of 0x80000000. For example, comparator 350 can also determine if adder 314 is above its half way mark based on the upper bits of adder 314. Consequently, if stat 312 is counting at 27 million counts/s, stat 312 will reach 0x80000000 in 80 seconds. If the stat 312 is not
- 10 polled or collected within 160 seconds (in approximately 3 minutes), stat 312 will roll over and the count value for stat 312 will become unreliable and inaccurate.

As shown in **Table 2** below, exemplary polling schemes and half count times (time an adder will reach 50% a maximum count for a 32-bit adder) for varying types of connections are illustrated.

15

Table 2
(Exemplary Polling Schemes)

Active Conns.	Average Cell Rate	Half Count Time	Average Polling Frequency
1	27000000	80 seconds	80 seconds/1 Conn = 80s
50	540000	66 minutes	66 mins/50 Conns = 80s
1000	27000	22 hours	22 hrs/1000 Conns = 80s
8000	3375	7.3 days	7.3 days/8000 Conns = 80s
256k	103	7.9 months	7.9 mins/256k Conns = 80s

As can be seen from **Table 2** above, to avoid losing reliable statistical information, statistical information for high traffic connections must be collected before collecting statistical information for low traffic connections. That is, if a connection, which is communicating high bandwidth traffic, is to be polled, e.g., at every 5 minutes, the polled information may be unreliable as a result of a stat rolling over in approximately 3 minutes. As such, a reliable way to poll or collect stats for a high bandwidth connection is to poll or collect all stats if any of its corresponding stats are greater than or equal to a threshold to ensure that its stats are collected before any of its counters roll over.

At operation 408, to collect a stat 312 that is greater than or equal to a threshold, FIFO control unit 308 stores its LCI in FIFO buffer 304.

Furthermore, FIFO buffer 304 may provide a plurality of entries such that a plurality of adders may reach a threshold thus causing a plurality of LCIs to be stored therein. FIFO control unit 308 may also prevent duplicate LCIs from being stored in FIFO buffer 304.

For one embodiment, if FIFO buffer 304 is full, FIFO control unit 308 may allow adder 314 to reach another threshold and place the LCI related to stat 312 in FIFO buffer 304. For example, if adder 314 goes from 0x80000000 to 0x80000001, FIFO control unit 308 may retry to load the LCI associated with adder 314 into FIFO buffer 304 a second time.

For an alternative embodiment, to address the situation of multiple counters reaching the threshold mark at the same time (e.g., halfway marks in

which multiple counters go from 0x7FFFFFFF to 0x80000000 simultaneously), a two-tier priority scheme may be implemented where the top bit of the address indicate that the LCI associated with the adder should be stored in into FIFO buffer 304 and the next to top bit of the adder may indicate:

- 5 • 0 - for low priority, LCI associated with the adder is put into FIFO buffer 304 if space is available.
- 1 - for high priority, LCI put into FIFO buffer 304 if space is available, if no space is available, then low priority entries may be removed and replaced with high priority entries.

10 As such, stats may reach, for example, a 75% maximum count value threshold. Such a priority scheme allows these stats to be serviced ahead of stats that are between one half and three quarters of their full count.

At operation 410, FIFO control unit 308 then sends an interrupt signal on interrupt signal line 306 to CPU 302. The interrupt signal informs CPU 302 that a
15 count value for adder 314 is greater than or equal to the threshold and the incremented stat is considered a "high priority" stat. This operation relates to operation 422 of **Figure 4B**. That is, after receiving an interrupt signal, CPU 320 will then read the LCI in FIFO buffer 304 (i.e., LCI₁) and poll or collect the high priority stat related to LCI₁. Operation 400 then loops back to operation 402 and
20 repeats the above operation.

Thus, a high priority interrupt operation for determining high priority and low priority stats has been described, which is used in polling and collecting stats.

Figure 4B is a flow chart illustrating one embodiment of an improved polling operation 420. For purposes of explanation, operation 420 refers to the exemplary stats collection architecture 206 of **Figure 3**. For purposes of explanation, operation 400 refers to a single adder 314, single stat 312, and a single comparator 350. Operation 400, however, may apply to a plurality of adders, stats, and comparators operating concurrently. Also, for purposes of explanation, operation 420 begins at operation 422.

Prior to operation 422, CPU 302 may poll each stat in stats memory 310 in regular time intervals. CPU 302, however, may poll and collect high priority stats before polling other low priority stats. Such an operation will now be described below.

Referring to **Figure 4B**, at operation 422, CPU 302 determines if there are any high priority stats. CPU 302 can determine if there is a high priority stat if CPU 302 receives an interrupt signal from FIFO control unit 308. That is, high priority stats determination is dependent on operation 410 of **Figure 4A** that causes an interrupt for when a count value for adder 314 is greater than or equal to a threshold count value and its associated stat is considered "high priority".

At operation 424, if CPU 302 does not determine that there is a high priority stat, CPU 302 will poll stats at regular intervals or use low priority LCIs

to poll a stat in stats memory 310. A low priority LCI relates to a connection having a count value below a threshold. For example, if the threshold is 50% of a maximum count value for an adder, CPU 302 can poll stats for connections in regular time intervals having a corresponding adder, which has a count value
5 below the threshold.

At operation 426, if there is a high priority stat (i.e., a stat related to an adder having a count value greater than or equal to a threshold), CPU 302 retrieves a LCI from FIFO buffer 304, which stores LCIs for high priority stats.

At operation 428, CPU 302 will then read or ("collect") all stats related to
10 the retrieved LCI from FIFO buffer 304, which is related to the high priority stat. CPU 302 will then store the collected stat in CPU memory 318. For example, if stat 312 was a high priority stat that caused an interrupt its corresponding LCI to be stored in FIFO buffer 304 as LCI₁, CPU 302 will read LCI₁ and retrieve next stat 312 corresponding to LCI₁. Operation 420 then loops back to operation 422 and
15 repeats the above operation.

Thus, an improved polling operation for collecting stats has been described by polling high priority stats before low priority stats.

Figure 5 is a flow chart illustrating one embodiment of an operation 500 for collecting statistics for the exemplary stats collection architecture 206 of
20 **Figure 3**. For purposes of explanation, operation 500 begins at operation 502.

At operation 502, statistics are counted for a plurality of connections. For example, a plurality of adders (i.e., adder 314) may count simultaneously statistics for a plurality of connections.

At operation 504, a determination is made to determine if a count value for each of the connections is greater than or equal to a threshold. For example, adder 314 may be a 32-bit counter and if its count value is equal to or above 0x80000000 it will store the LCI for that stat (i.e., stat 312) associated with adder 314 in FIFO buffer 304. If no adder is equal to or above the threshold, operation 500 continues at operation 502.

At operation 506, if a counter is greater than or equal a threshold, CPU 302 will collect statistics for the connections associated with counters having a count value greater than or equal to the threshold before collecting statistics for connections having a count value below the threshold. For example, CPU 302 will be interrupted to collect stats based on the LCIs stored in FIFO buffer 304. Thus, CPU will collect statistics associated with counters before they roll over to prevent unreliable and inaccurate stats information to be stored in stats memory 310. Operation 500 then loops back to operation 502 and repeats the above operation.

Thus, an operation for collecting statistics that collects reliable and accurate stats has been described.

In the foregoing specification, the invention has been described with reference to specific exemplary embodiments thereof. It will, however, be

evident that various modifications and changes may be made thereto without departing from the broader spirit and scope of the invention as set forth in the appended claims. The specification and drawings are, accordingly, to be regarded in an illustrative sense rather than a restrictive sense.

1. a) $\frac{1}{2} \log_2 \frac{1}{2}$	
1.000000	0.500000
0.999999	0.499999
0.999998	0.499998
0.999997	0.499997
0.999996	0.499996
0.999995	0.499995
0.999994	0.499994
0.999993	0.499993
0.999992	0.499992
0.999991	0.499991
0.999990	0.499990
0.999989	0.499989
0.999988	0.499988
0.999987	0.499987
0.999986	0.499986
0.999985	0.499985
0.999984	0.499984
0.999983	0.499983
0.999982	0.499982
0.999981	0.499981
0.999980	0.499980
0.999979	0.499979
0.999978	0.499978
0.999977	0.499977
0.999976	0.499976
0.999975	0.499975
0.999974	0.499974
0.999973	0.499973
0.999972	0.499972
0.999971	0.499971
0.999970	0.499970
0.999969	0.499969
0.999968	0.499968
0.999967	0.499967
0.999966	0.499966
0.999965	0.499965
0.999964	0.499964
0.999963	0.499963
0.999962	0.499962
0.999961	0.499961
0.999960	0.499960
0.999959	0.499959
0.999958	0.499958
0.999957	0.499957
0.999956	0.499956
0.999955	0.499955
0.999954	0.499954
0.999953	0.499953
0.999952	0.499952
0.999951	0.499951
0.999950	0.499950
0.999949	0.499949
0.999948	0.499948
0.999947	0.499947
0.999946	0.499946
0.999945	0.499945
0.999944	0.499944
0.999943	0.499943
0.999942	0.499942
0.999941	0.499941
0.999940	0.499940
0.999939	0.499939
0.999938	0.499938
0.999937	0.499937
0.999936	0.499936
0.999935	0.499935
0.999934	0.499934
0.999933	0.499933
0.999932	0.499932
0.999931	0.499931
0.999930	0.499930
0.999929	0.499929
0.999928	0.499928
0.999927	0.499927
0.999926	0.499926
0.999925	0.499925
0.999924	0.499924
0.999923	0.499923
0.999922	0.499922
0.999921	0.499921
0.999920	0.499920
0.999919	0.499919
0.999918	0.499918
0.999917	0.499917
0.999916	0.499916
0.999915	0.499915
0.999914	0.499914
0.999913	0.499913
0.999912	0.499912
0.999911	0.499911
0.999910	0.499910
0.999909	0.499909
0.999908	0.499908
0.999907	0.499907
0.999906	0.499906
0.999905	0.499905
0.999904	0.499904
0.999903	0.499903
0.999902	0.499902
0.999901	0.499901
0.999900	0.499900
0.999899	0.499899
0.999898	0.499898
0.999897	0.499897
0.999896	0.499896
0.999895	0.499895
0.999894	0.499894
0.999893	0.499893
0.999892	0.499892
0.999891	0.499891
0.999890	0.499890
0.999889	0.499889
0.999888	0.499888
0.999887	0.499887
0.999886	0.499886
0.999885	0.499885
0.999884	0.499884
0.999883	0.499883
0.999882	0.499882
0.999881	0.499881
0.999880	0.499880
0.999879	0.499879
0.999878	0.499878
0.999877	0.499877
0.999876	0.499876
0.999875	0.499875
0.999874	0.499874
0.999873	0.499873
0.999872	0.499872
0.999871	0.499871
0.999	

We claim:

- 1 1. A method for a network switch, the method comprising:
2 counting statistics for a plurality of connections;
3 determining if a count value for each of the connections is above a
4 threshold; and
5 collecting statistics for the connections having a count value above the
6 threshold before connections having a count value below the threshold.
- 1 2. The method of claim 1, wherein the threshold is a percentage of a
2 maximum count value.
- 1 3. The method of claim 1, wherein determining if the count value for each of
2 the connections is above the threshold includes:
3 sending an interrupt signal for each connection having a determined
4 count value above the threshold.
- 1 4. The method of claim 3, wherein collecting statistics for the connections
2 includes:
3 collecting statistics for connections with corresponding interrupt signals
4 before other connections with no corresponding interrupt signals.

1 5. The method of claim 3, wherein sending an interrupt signal includes:
2 placing a logical connection identifier for each statistic having an
3 associated count value above the threshold in a first-in-first-out (FIFO) buffer.

1 6. The method of claim 5, wherein collecting statistics for connections
2 includes:
3 reading statistics in a memory based on the logical connection identifiers
4 stored in the FIFO buffer.

1 7. An apparatus comprising:
2 one or more counters, each counter to count statistics for a connection; and
3 a processor to determine if a count value for each of the counters is above
4 a threshold and to collect statistics from counters having a count value above the
5 threshold before counters having a count value below the threshold.

1 8. The apparatus of claim 7, wherein the threshold is a percentage of a
2 maximum count value.

1 9. The apparatus of claim 7, wherein each counter is to send an interrupt
2 signal for a respective connection having a determined count value above the
3 threshold.

1 10. The apparatus of claim 9, wherein the processor is to collect statistics for
2 counters having corresponding interrupt signals before other counters with no
3 corresponding interrupt signals.

1 11. The apparatus of claim 9, further comprising:
2 a first-in-first-out (FIFO) buffer to store logical connection identifiers, each
3 logical connection identifier to be used to collect high priority statistics.

1 12. The apparatus of claim 11, further comprising:
2 a memory to store statistics, wherein the processor is to read high priority
3 statistics in the memory based on the logical connection identifiers stored in the
4 FIFO buffer.

1 13. A network interface comprising:
2 a memory to store statistics for a connection;
3 one or more counters, each counter to count statistics stored in the
4 memory;
5 a processor to determine if a count value for each of the counters is above
6 a threshold and to collect statistics stored in the memory derived from counters
7 having a count value above the threshold before collecting statistics in the
8 memory derived from counters having a count value below the threshold.

1 14. The network interface of claim 13, wherein the threshold is a percentage of
2 a maximum count value.

1 15. The network interface of claim 13, wherein each counter is to send an
2 interrupt signal for a respective connection having a determined count value
3 above the threshold.

1 16. The network interface claim 15, wherein the processor is to collect
2 statistics for counters with corresponding sent interrupt signals before other
3 counters with no interrupt signals.

1 17. The network interface of claim 15, further comprising:
2 a first-in-first-out (FIFO) buffer to store logical connection identifiers, each
3 logical connection identifier to be used to collect high priority statistics.

1 18. The network interface of claim 17, further comprising:
2 a memory to store statistics, wherein the processor is to read high priority
3 statistics in the memory based on the logical connection identifiers in the FIFO
4 buffer.

1 24. The apparatus of claim 23, wherein the means for collecting statistics
2 collects statistics for connections with corresponding interrupt signals before
3 other connections with no corresponding interrupt signals.

1 25. The apparatus of claim 23, further comprising:
2 means for placing a logical connection identifier for each statistic having
3 an associated count value above the threshold in a first-in-first-out (FIFO) buffer.

1 26. The apparatus of claim 25,
2 means for reading statistics in a memory based on the logical connection
3 identifiers stored in the FIFO buffer.

1 27. A statistics collection module comprising:
2 a memory to store statistics for a connection;
3 one or more module counters, each module counter to count statistics
4 stored in the memory;
5 a central processing unit (CPU) having internal CPU counters to count
6 values from the one or more module counters, the CPU to determine if a count
7 value for each of the module counters is greater than or equal to a threshold and
8 to collect statistics stored in the memory derived from the module counters
9 having a count value that is greater than or equal to the threshold before

10 collecting statistics in the memory derived from the module counters having a
11 count value below the threshold.

1 28. The statistics collection module of claim 27, wherein the internal CPU
2 counters are wider than the one or more module counters.

1 29. The statistics collection module of claim 28, wherein the internal CPU
2 counters are 64-bit counters and the one or more module counters are 32-bit
3 counters.

ABSTRACT

An apparatus and method that provides improved statistics collection for high bandwidth interfaces supporting multiple connections. Statistics are counted for a plurality of connections. A determination is made to determine if a count value for each of the connections is greater than or equal to a threshold.

- 5 Statistics are collected for connections having a determined count value greater than or equal to the threshold before connections having a count value below the threshold. The apparatus and method can thus avoid data loss and reduce wasted processing time in maintaining reliable and accurate statistics for multiple connection supported by a network switch or interface.

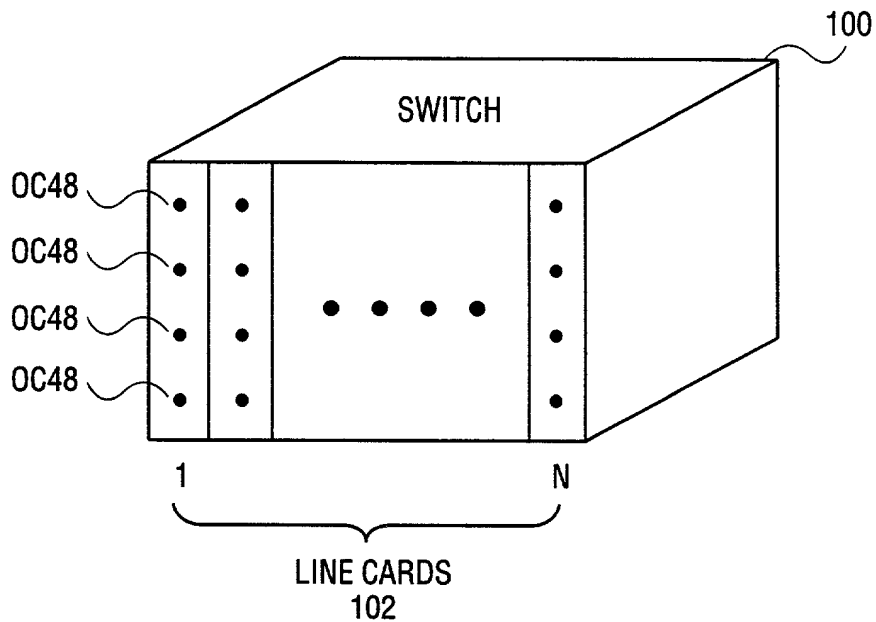


FIG. 1

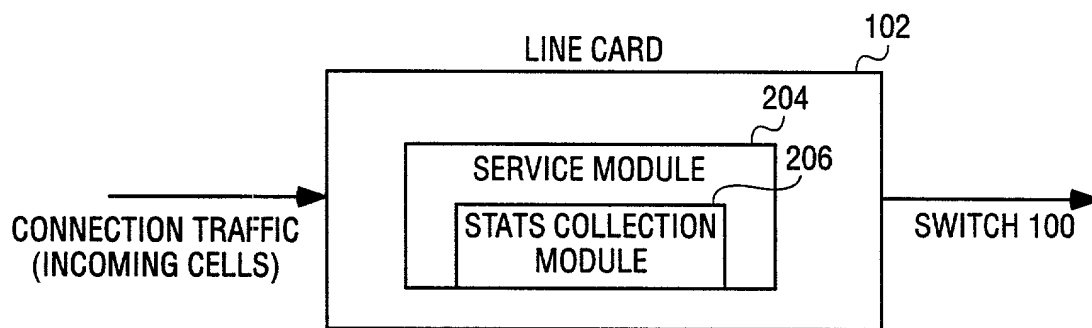


FIG. 2

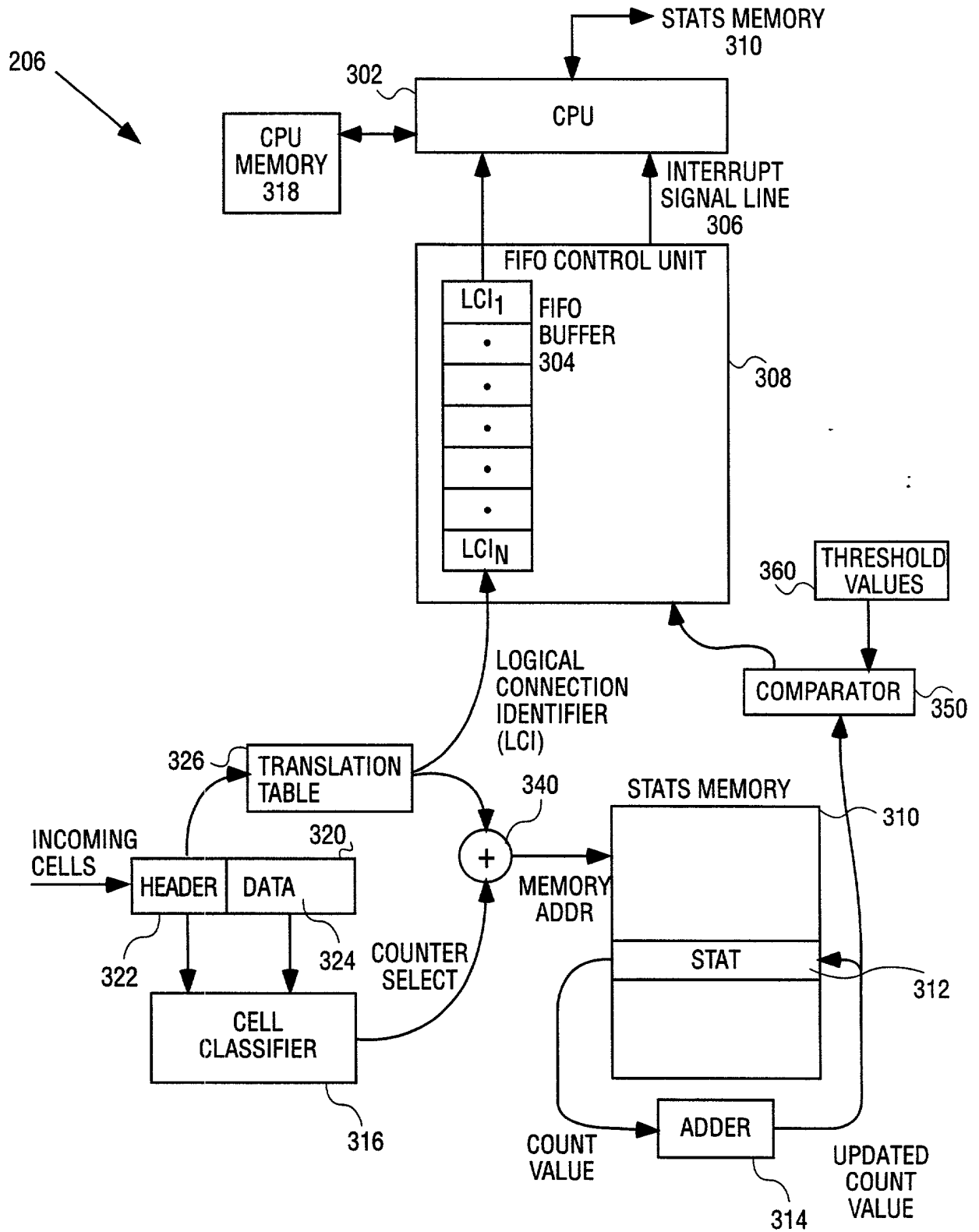


FIG. 3

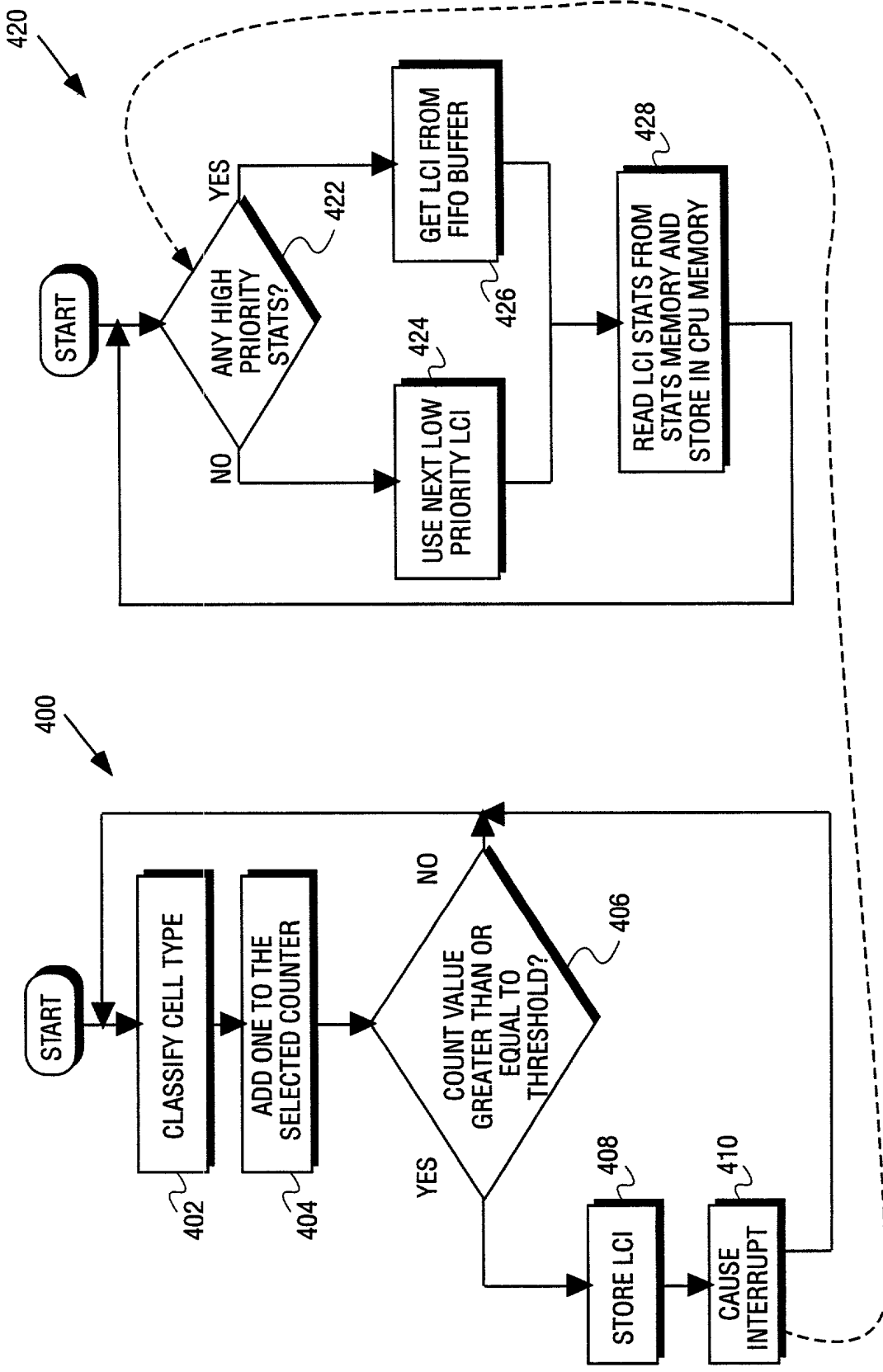


FIG. 4B

FIG. 4A

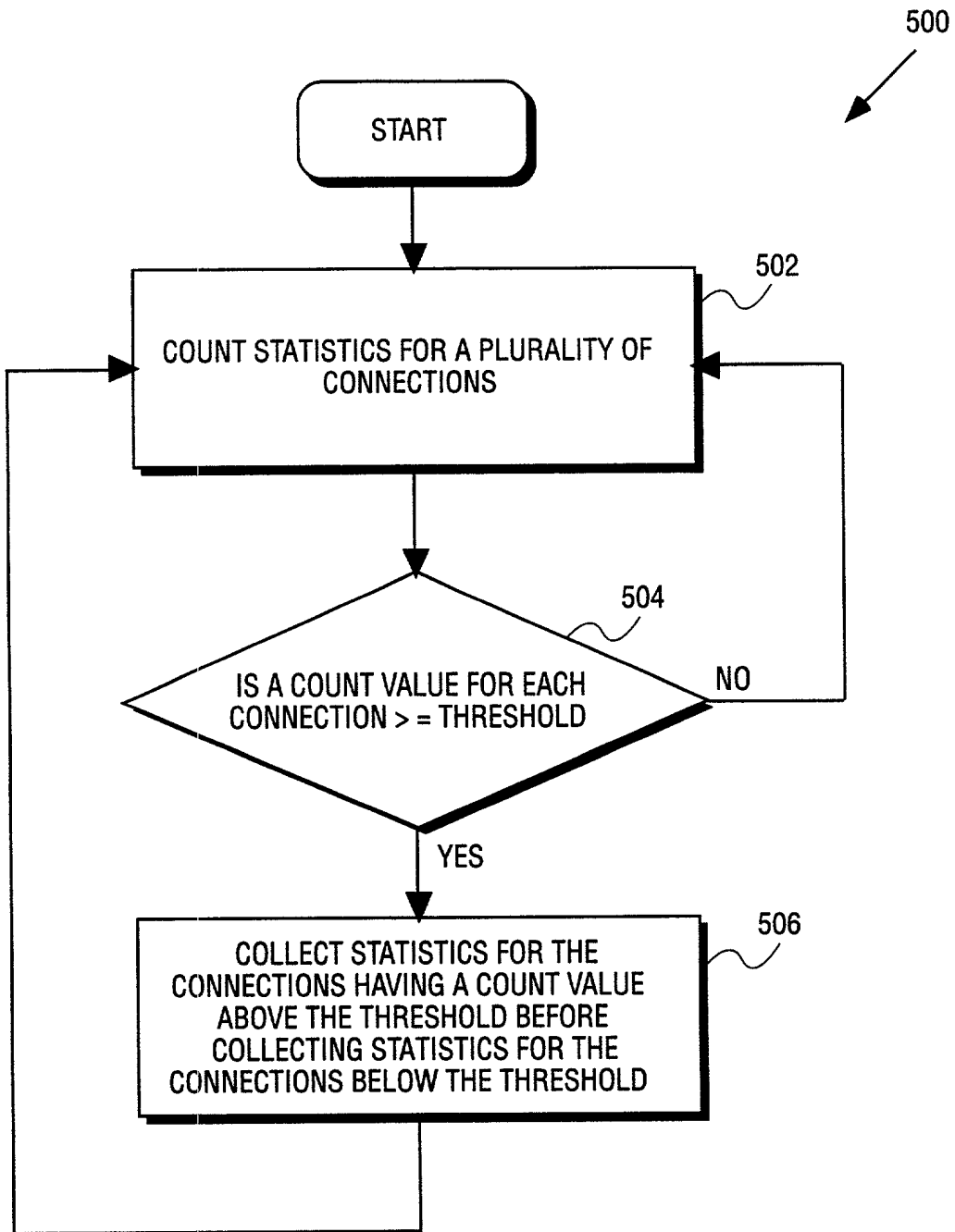


FIG. 5